Fifty Years of Computing at LLNL as a Lens to the Future

ISC 2002

June 22, 2002

Dona L. Crawford

Computation Associate Director

Lawrence Livermore National Laboratory

UCRL-PRES-149571

Contributors

- Tom Adams
- Alane Alchorn
- Bernie Alder
- Peter Brown
- Bill Camp, SNL
- Randy Christensen
- Evi Dube
- Jim Frank
- Kent Johnson
- David Keyes
- Bill Lokke
- Mike McCoy
- Jim McGraw
- Tom Manteuffel, U. Colorado
- John May
- Hans Meuer, U. Mannheim
- George Michael
- Pat Miller
- Cynthia Nitta
- Joe Requa
- Mark Seager
- Kathy Turnbeaugh
- Mike Vahle, SNL
- Dick Watson
- Charlie Westbrook
- Bing Young
- George Zimmerman

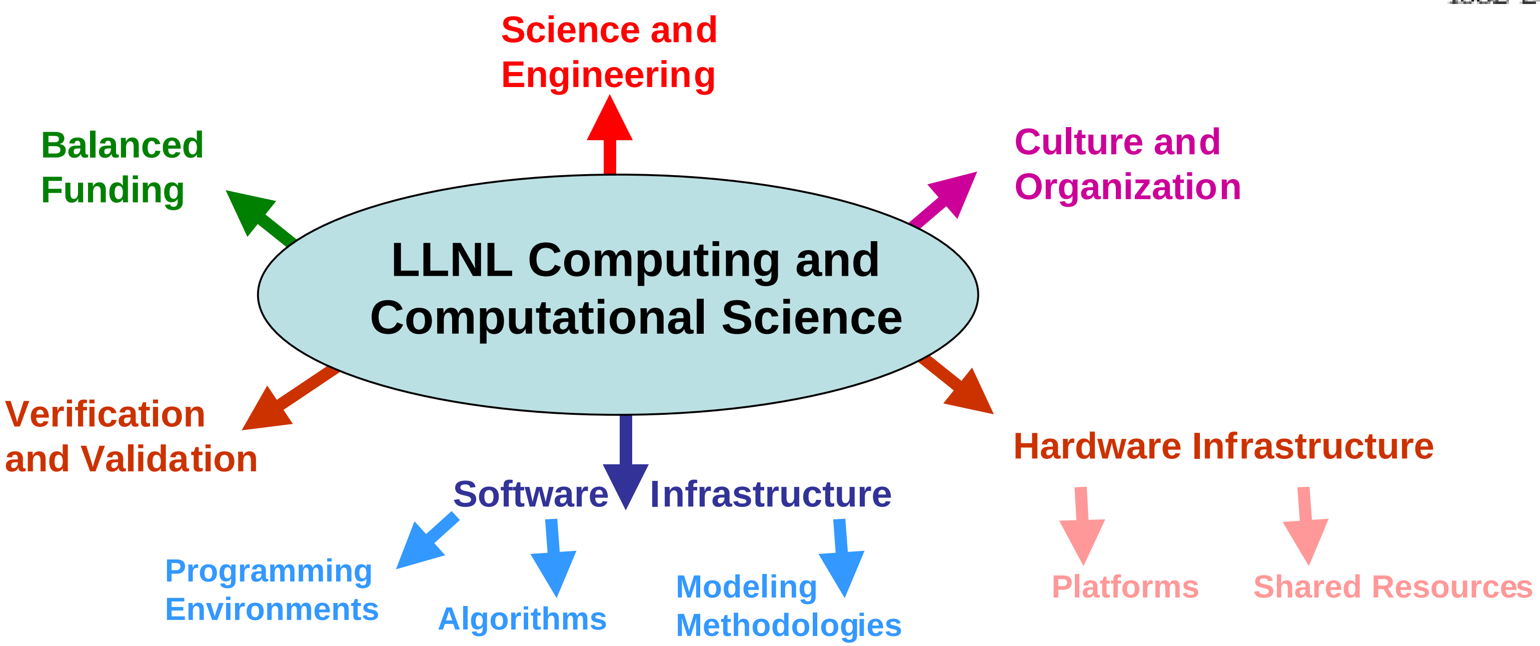
“We know what we are, but know not what we may be.”

William Shakespeare, 1694

We may not know the future in detail, but certain issues existed 50 years ago; those issues persist today and will accompany us into the future.

Computing & computational science

the past and the present inform the future



While the future will hold amazing advances, key issues from the past and the present along each dimension noted above will accompany high-performance computing into the future

- Balance is crucial within and across these dimensions

Computational science exploits recognized scientific practices

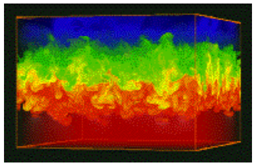
How does science work? (There can be no single ***right*** way to do science, but some threads endure.)

- Exemplars and organization
  - combinations of theory, research agendas, practices and methodologies
  - social structures and culture (teams, training, rewards, texts)
- Evidence (repeatable experiments and observations)
- Communal openness to confirmation or rejection

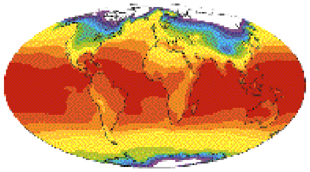
Computational modeling has brought changes in methodology and emphasis to all three

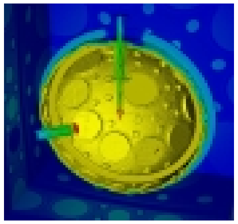
The desire to understand complex physical phenomena drives computing and computational science

Lawrence Livermore National Laboratory
Making History
Making a Difference
1952-2002



Experiments are prohibited





SSP / ASCI
radiation transport
hydrodynamics

Environmental
global climate
groundwater flow

Engineering
structural dynamics
electromagnetics

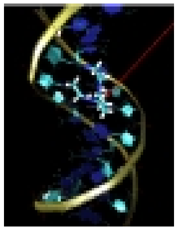
Experiments are impractical

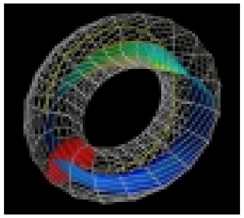
Experiments are too expensive

Physics & Biology
materials modeling
drug design

Enabling Computational Technologies
computational math
computer science
MPP

Lasers & Energy
combustion
ICF modeling





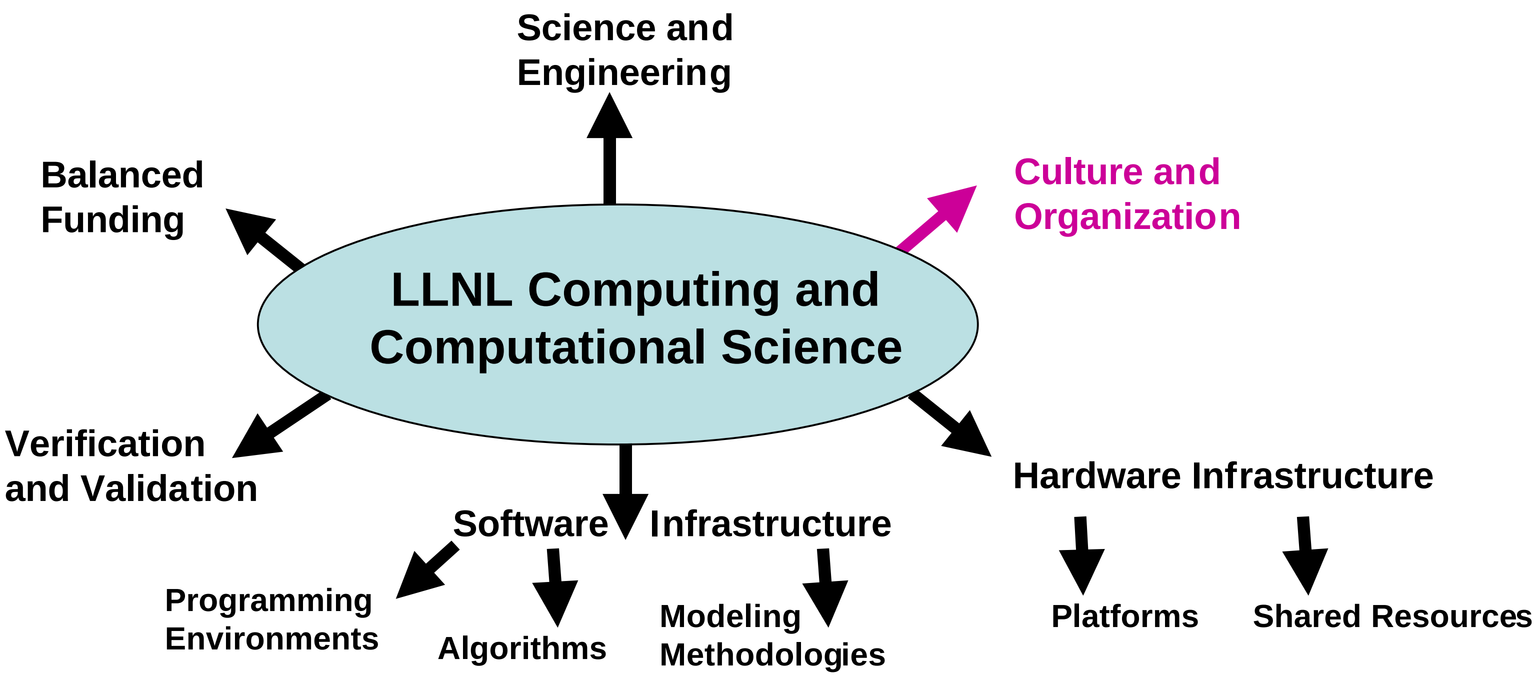
Simulation challenges theory and experiment to refine quality and accuracy

The expected applications will continue to need more capability

Lawrence Livermore National Laboratory
Making History Making a Difference
1952-2002

- biology, biochemistry, biomedicine, molecular nano technology
- scientific and business data mining
- web searching
- national scale economic modeling
- chemistry, chemical engineering
- physics
- space science and astronomy
- artificial intelligence
- climate and weather studies
- environmental studies
- geophysics and petroleum engineering
- aerospace, mechanical, and manufacturing engineering
- military applications
- nuclear weapons stewardship
- cryptography and digital signal processing
- multiuser immersive environments
- symbolic and experimental mathematics
- business operations
- general societal problems
- disaster modeling

Lab culture and organization nurture computing excellence



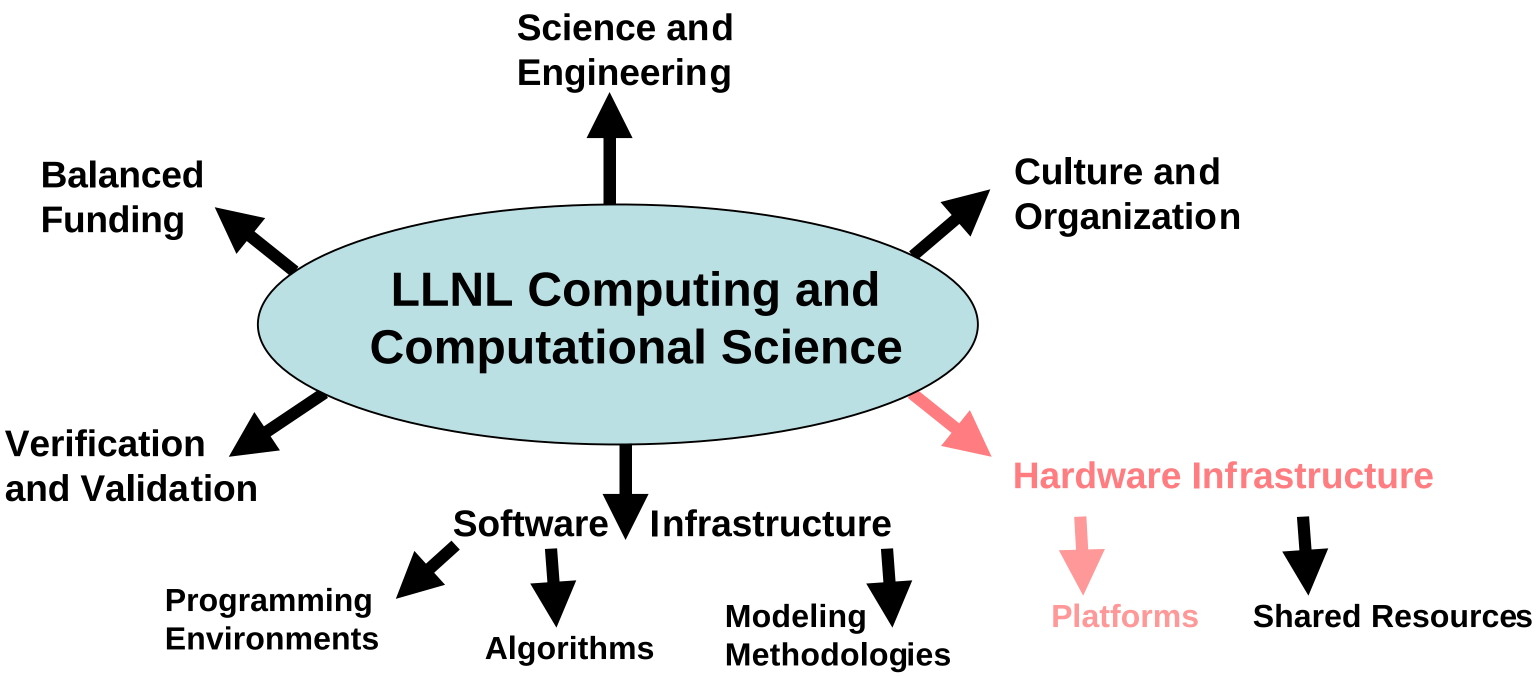
Unique culture and organization shape LLNL computing

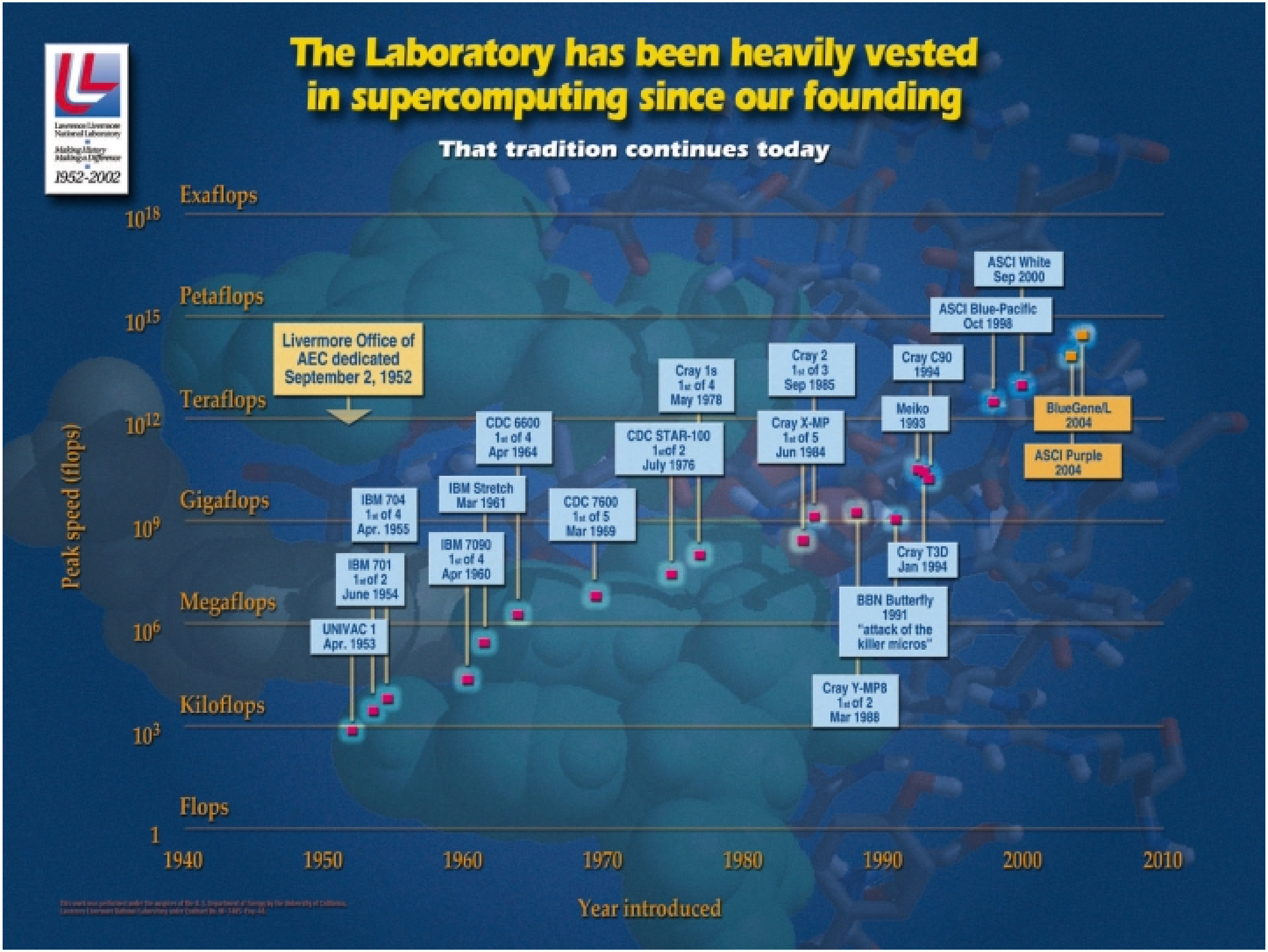
- Past: LLNL founders recognized the importance of coupling theory to experimental and computational capabilities
- Present and future: Lessons learned early and reinforced over time create today's culture and organization — and that needed for the future
  - Belief in the value of computational modeling
  - Teams are key
  - Individual contributors are highly valued
  - Visionary, enthusiastic, dedicated (some say fanatical) users are critical
  - Teams work in a milieu of various model styles
  - Collaborate freely and publish widely

Edward Teller, LLNL co-founder

Lawrence Livermore National Laboratory
Making History Making a Difference
1952-2002

Supercomputing platforms at LLNL





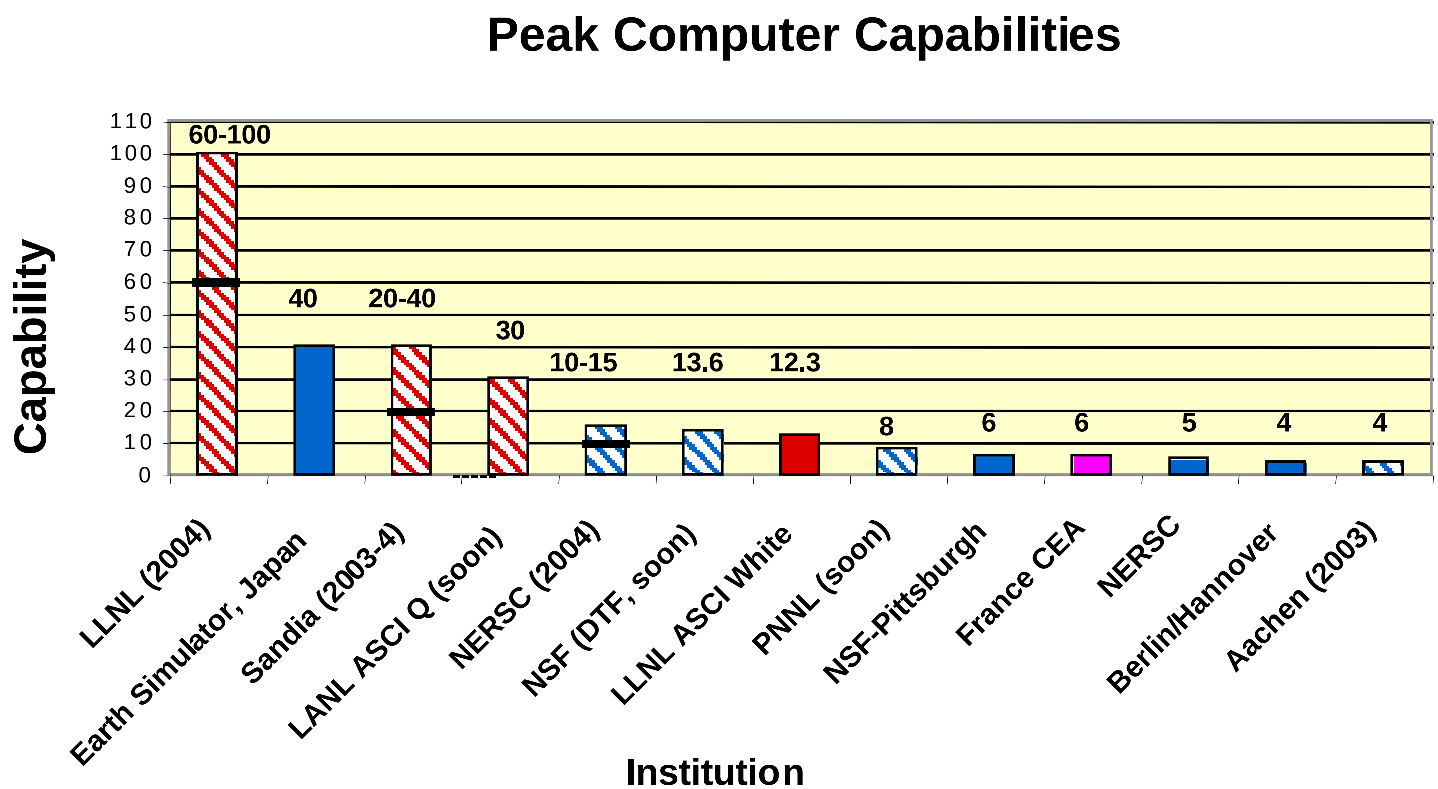
The Laboratory has been heavily vested in supercomputing since our founding
That tradition continues today
Lawrence Livermore National Laboratory
Making History Making a Difference
1952-2002
Exaflops
Petaflops
Teraflops
Gigaflops
Megaflops
Kiloflops
Flops
10^{18}
10^{15}
10^{12}
10^{9}
10^{6}
10^{3}
1
Peak speed (flops)
1940
1950
1960
1970
1980
1990
2000
2010
Year introduced
Livermore Office of AEC dedicated September 2, 1952
UNIVAC 1 Apr. 1953
IBM 701 1st of 2 June 1954
IBM 704 1st of 4 Apr. 1955
IBM 7090 1st of 4 Apr 1960
IBM Stretch Mar 1961
CDC 6600 1st of 4 Apr 1964
CDC 7600 1st of 5 Mar 1969
CDC STAR-100 1st of 2 July 1976
Cray 1s 1st of 4 May 1978
Cray X-MP 1st of 5 Jun 1984
Cray 2 1st of 3 Sep 1985
Cray Y-MP8 1st of 2 Mar 1988
BBN Butterfly 1991 "attack of the killer micros"
Cray T3D Jan 1994
Meiko 1993
Cray C90 1994
ASCI Blue-Pacific Oct 1998
ASCI White Sep 2000
BlueGene/L 2004
ASCI Purple 2004

Workhorses in the LLNL supercomputing stable

Lawrence Livermore National Laboratory
Making History Making a Difference
1952-2002

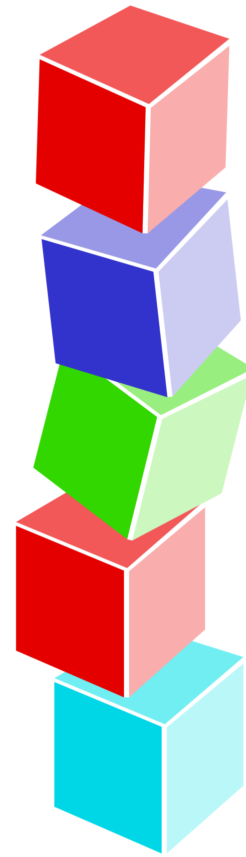
System	Manufacturer Model	Interconnect	Nodes	CPUs	Memory (GB)	Peak GFLOPS
Unclassified Network						
ASCI Frost	IBM SP	Colony DS	68	1,088	1,088	1,632
ASCI Blue-Pacific CTR	IBM SP	TB3	280	1,120	410	744
GPS Cluster	Compaq GS320/ES45	N/A	17	96	264	192
Tera Cluster 2000	Compaq SC ES40	Elan3	128	512	280	683
Linux Cluster	Compaq DS20E/API UP2000	N/A	38	76	76	101
Tera Cluster 98	Compaq 4100/DS20E/ES40	N/A	44	152	68	169
UOCF	Sun E6000	N/A	1	24	16	12
Viz Engine (Riptide)	SGI Onyx2	8 IR2 Pipes	1	48	16	24
Total						**3,557**
Classified Network						
ASCI White	IBM SP	Colony DS	512	8,192	6,256	12,288
ASCI Ice	IBM SP	Colony DS	28	448	448	672
ASCI Blue-Pacific SST	IBM SP	TB3/HPGN	1464	5,856	2,628	3,888
Sector S		TB3	488	1,952	1,164	1,296
Sector K		TB3	488	1,952	732	1,296
Sector Y		TB3	488	1,952	732	1,296
PCRA (Adelie)	Linux NetworX	Elan3	128	256	256	870
PCRB (Emperor)	Linux NetworX	Elan3	88	176	176	598
Furnace Cluster	API CS20	N/A	64	128	128	213
SC Cluster	Compaq ES40/ES45	N/A	40	160	384	235
ICF Cluster	Compaq ES40/DS10L	N/A	12	36	12	48
Forest Cluster	Digital 8400	MC 1.2	6	54	56	66
Tri-Lab Viz Engine (Whitecap)	SGI Onyx3800	4 IR3 Pipes	1	96	96	77
Viz Engine (Tidalwave)	SGI Onyx2	16 IR2 Pipes	1	64	24	38
Viz Engine (Edgewater)	SGI Onyx2	10 IR2 Pipes	1	40	18	24
Total						**19,017**

Terascale simulation environments are critical at many other institutions



The platform balancing act is a delicate one

- Supercomputing platforms must balance
  - Power
  - Memory size
  - Bandwidth
  - Latency
  - I/O
- Supercomputers after the Cray 1 began to lose balance



Applications require balanced systems

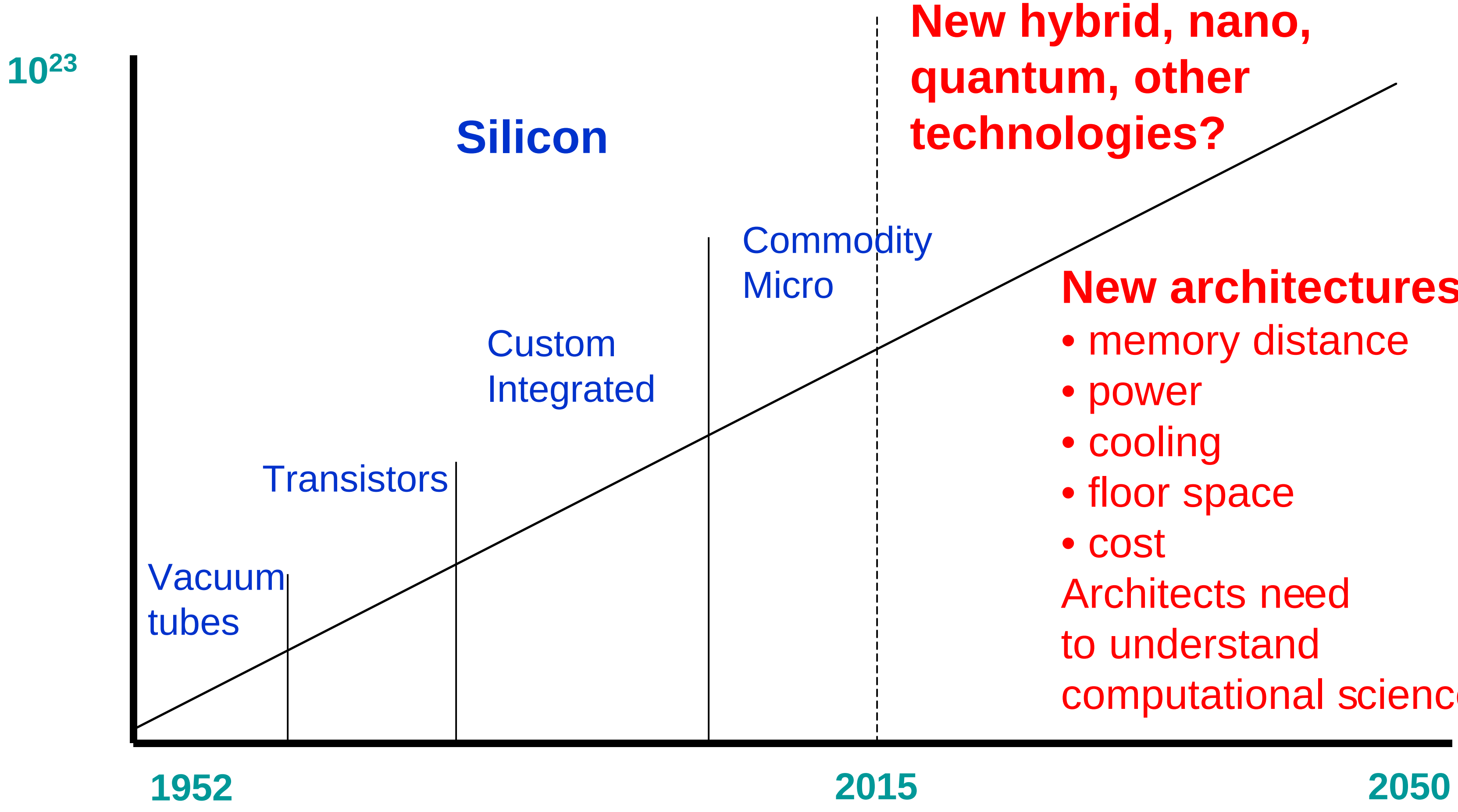
For every 1 gigaflop peak performance, we need

Capability (Flops)	***1990 (10^9)***	***2000 (10^{13})***	***2020 (10^{17})***	***2050 (10^{23})***
1 GB memory size	10^9	10^{13} *(10^{12})*	10^{17}	10^{23}
50 GB disk storage	5×10^{10}	5×10^{14} *(10^{14})*	5×10^{18}	5×10^{24}
10 TB archival storage	10^{13}	10^{17} *(10^{14})*	10^{21}	10^{27}
16 GB/s cache bandwidth	1.6×10^{10}	1.6×10^{14} *(10^{11})*	1.6×10^{18}	1.6×10^{24}
3 GB/s memory bandwidth	3×10^9	3×10^{13} *(10^{10})*	3×10^{17}	3×10^{23}
0.1 GB/s I/O bandwidth	10^8	10^{12} *(10^{10})*	10^{16}	10^{22}
0.01 GB/s disk bandwidth	10^7	10^{11} *(10^9)*	10^{15}	10^{21}
1 MB/s archival storage band.	10^6	10^{10} *(10^8)*	10^{14}	10^{20}

Where 10^6 = mega, 10^9 = giga, 10^{12} = tera, 10^{15} = peta, 10^{18} = exa, 10^{21} = zetta, and 10^{24} = yotta

Lawrence Livermore National Laboratory
Making History Making a Difference
1952-2002

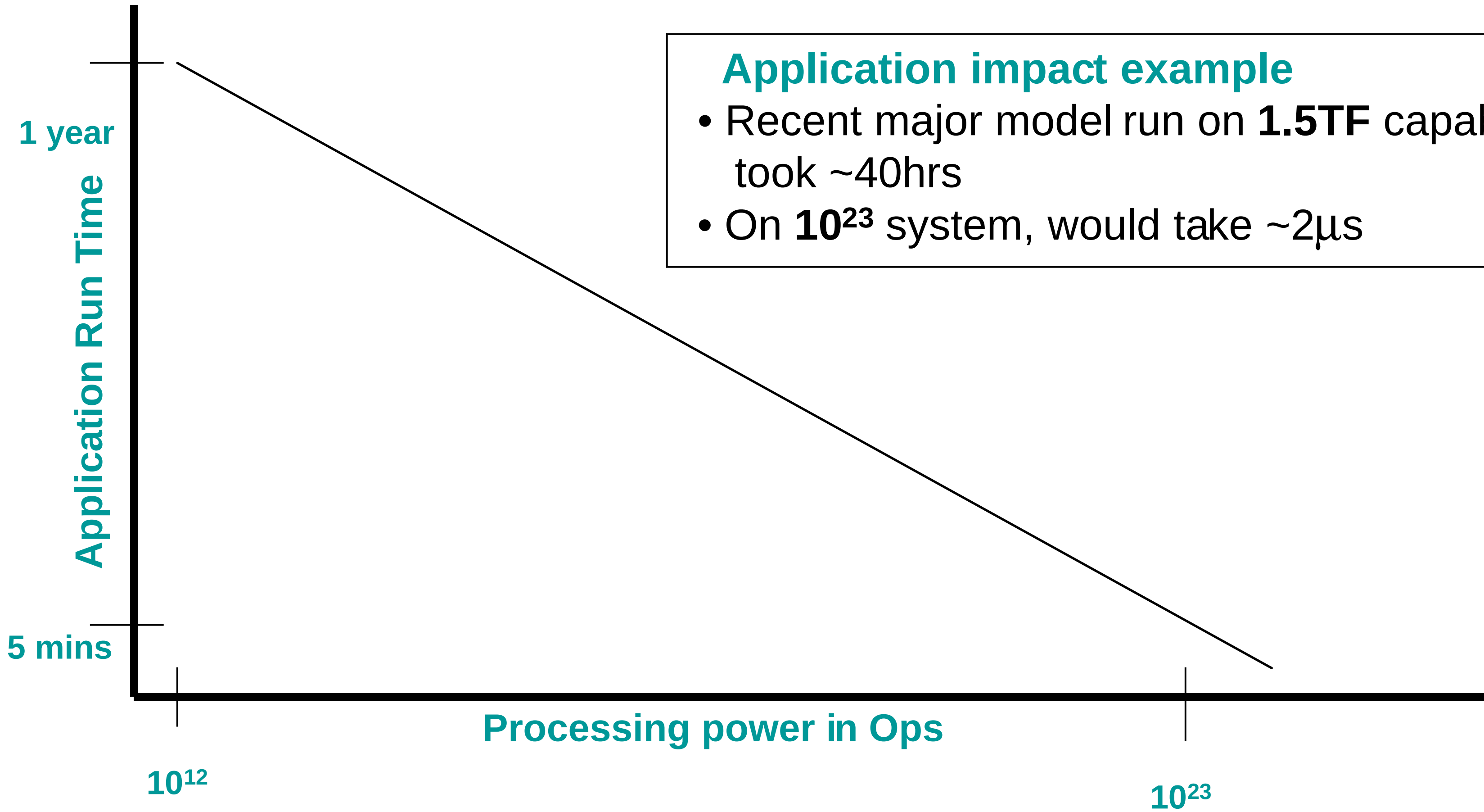
The challenge of maintaining growth in processing power



New architectures
- memory distance
- power
- cooling
- floor space
- cost

Architects need to understand computational science

Potential runtime improvements



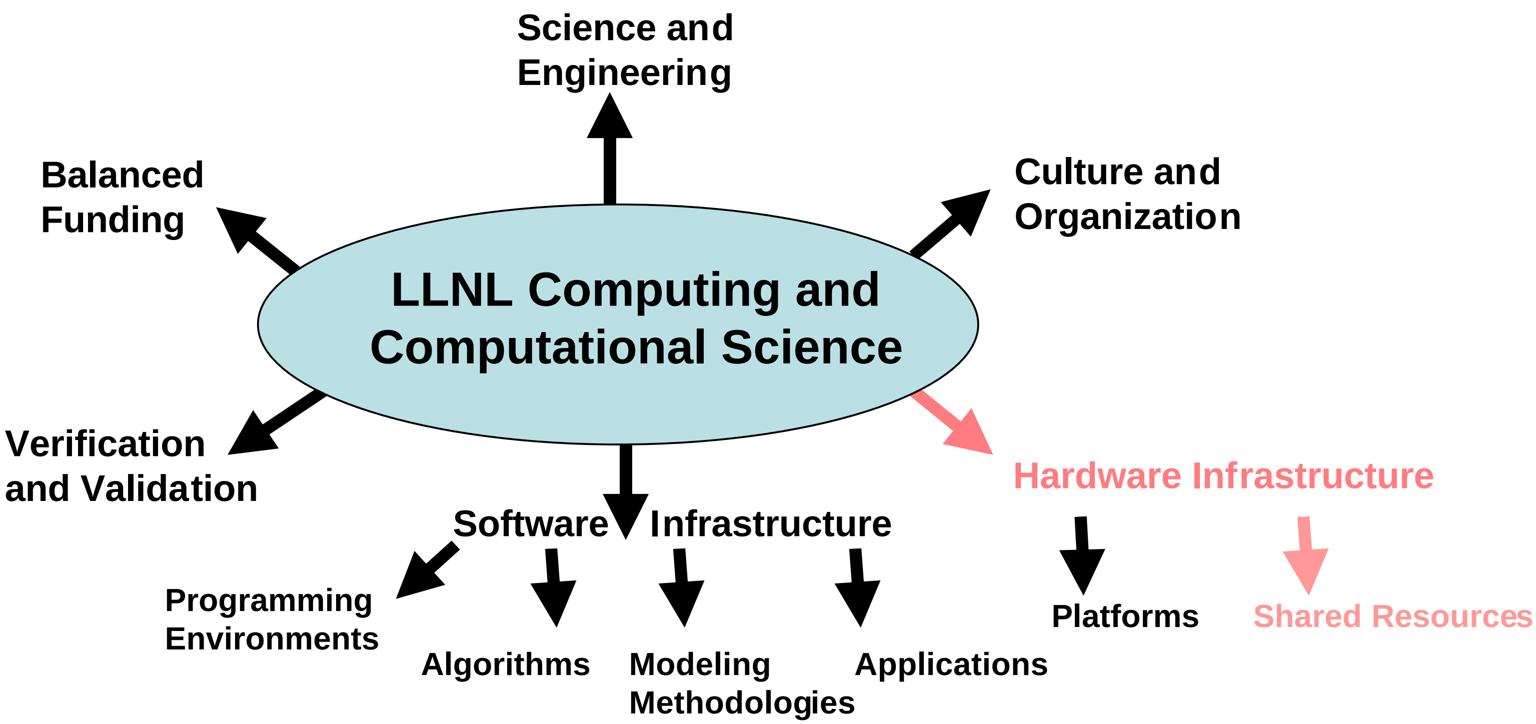
Application impact example

- Recent major model run on **1.5TF** capability took ~40hrs
- On **10^{23}** system, would take ~2μs

Vendor partnerships help accelerate capabilities

- Past systems
  - R&D funding to develop prototype systems (e.g. IBM Stretch, IBM Photostore)
  - Commit to buy first or early production systems
    (e.g. CDC 6600, 7600, Star, Cray 1, XMP, YMP, Cray 2, BBN Butterfly, Meiko)
- Present systems
  - ASCI Pathforward strategy - interconnects, rendering, storage
  - Largest machines such as ASCI Blue-Pacific, ASCI White
- Future systems
  - IBM Blue Gene-L
  - ASCI Purple and beyond

Shared resources at LLNL

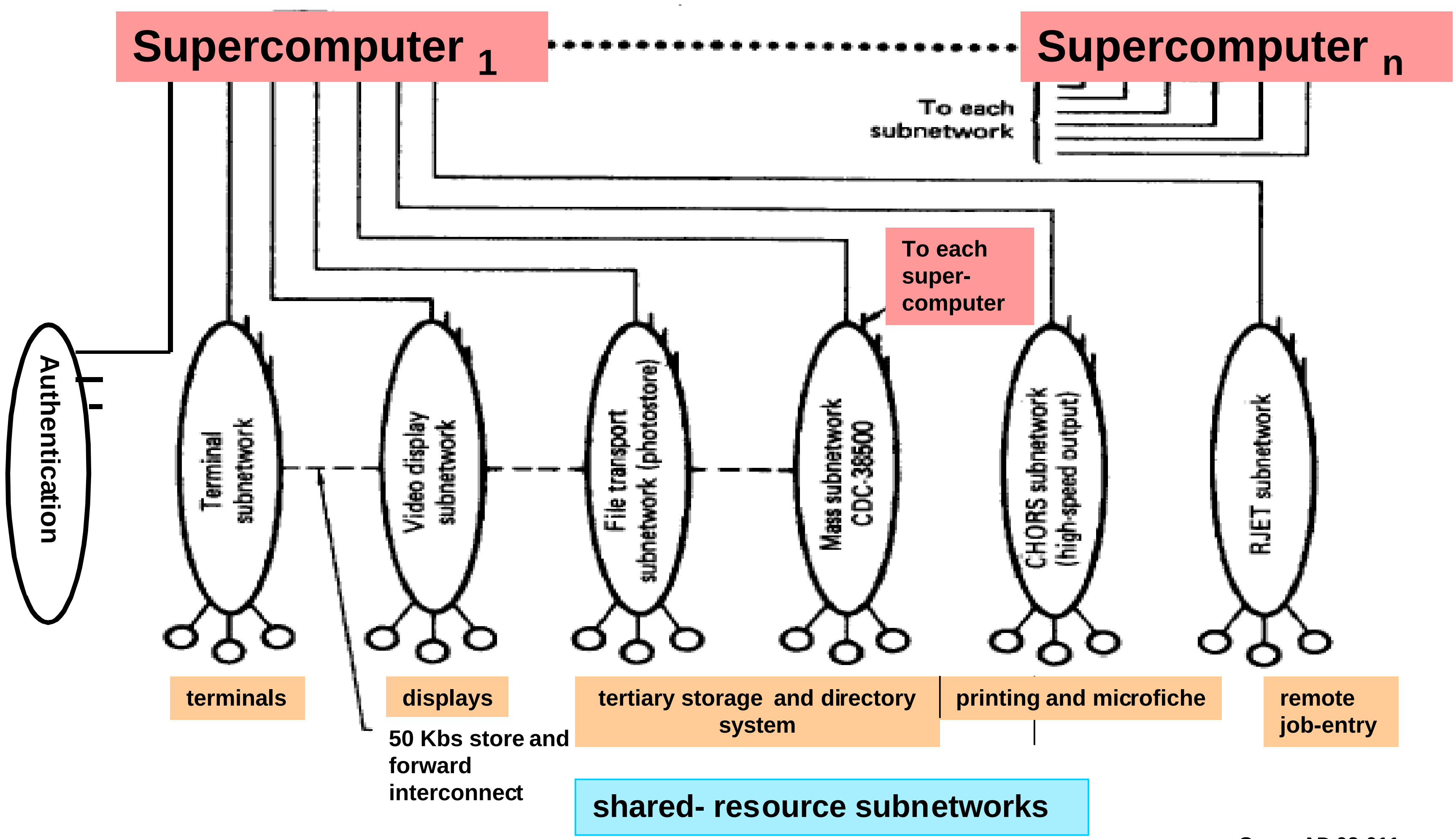


Balanced network-based shared-resource architecture

Lawrence Livermore National Laboratory
Making History Making a Difference
1952-2002

- High-performance computers are useless without the balanced environment of storage, networks, I/O, and security
- The need for fast, secure, network-centered, shared-resource architectures was recognized in the Sixties
- Getting funding agency to appreciate ongoing need for balanced shared resources to get full value from processing system is a continuing educational issue
  - Balanced system funding will continue as a future issue

The Octopus network design dominated the Seventies



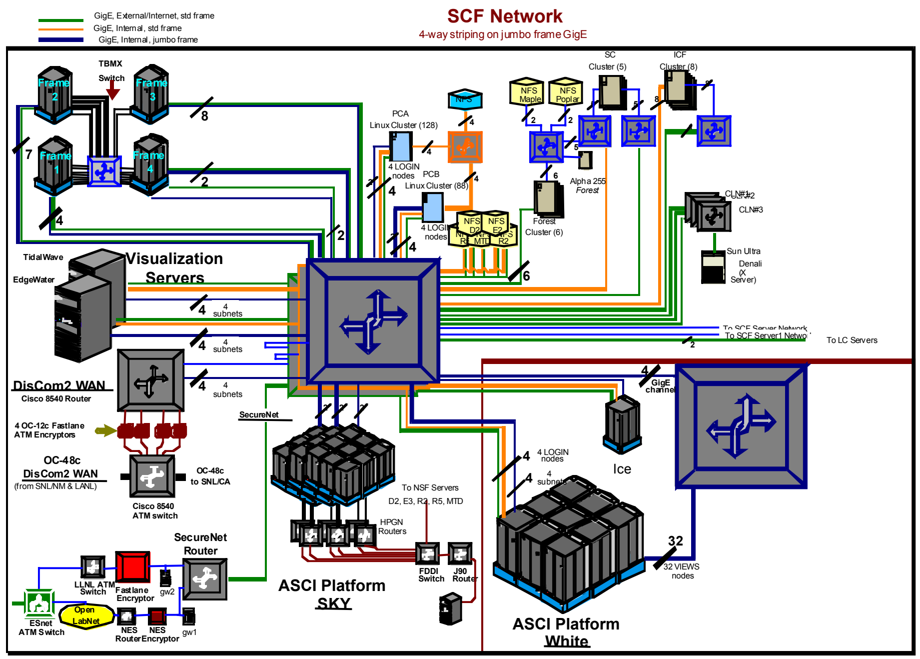
Today's shared-resource environment

Present

- -Gigabit Ethernet everywhere
  - 4 jumbo frame interfaces per machine
- -DisCom WAN
  - 2.4Gbps to SNL & LANL
  - >100MB/s FTP between sites

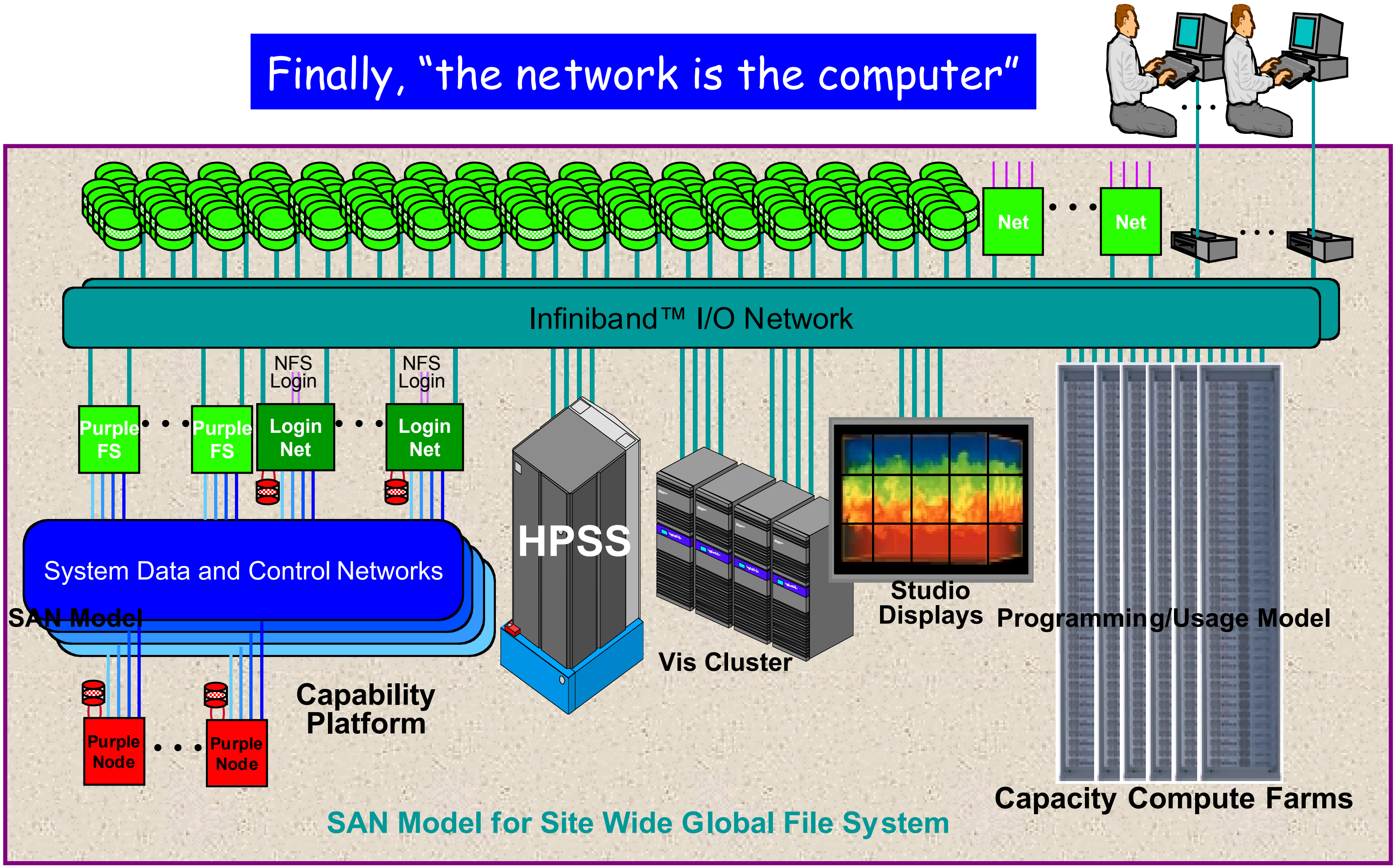
Near Future

- -10GigE between buildings
- -TCP Offload Engines (TOEs)



Future network environments will share all storage

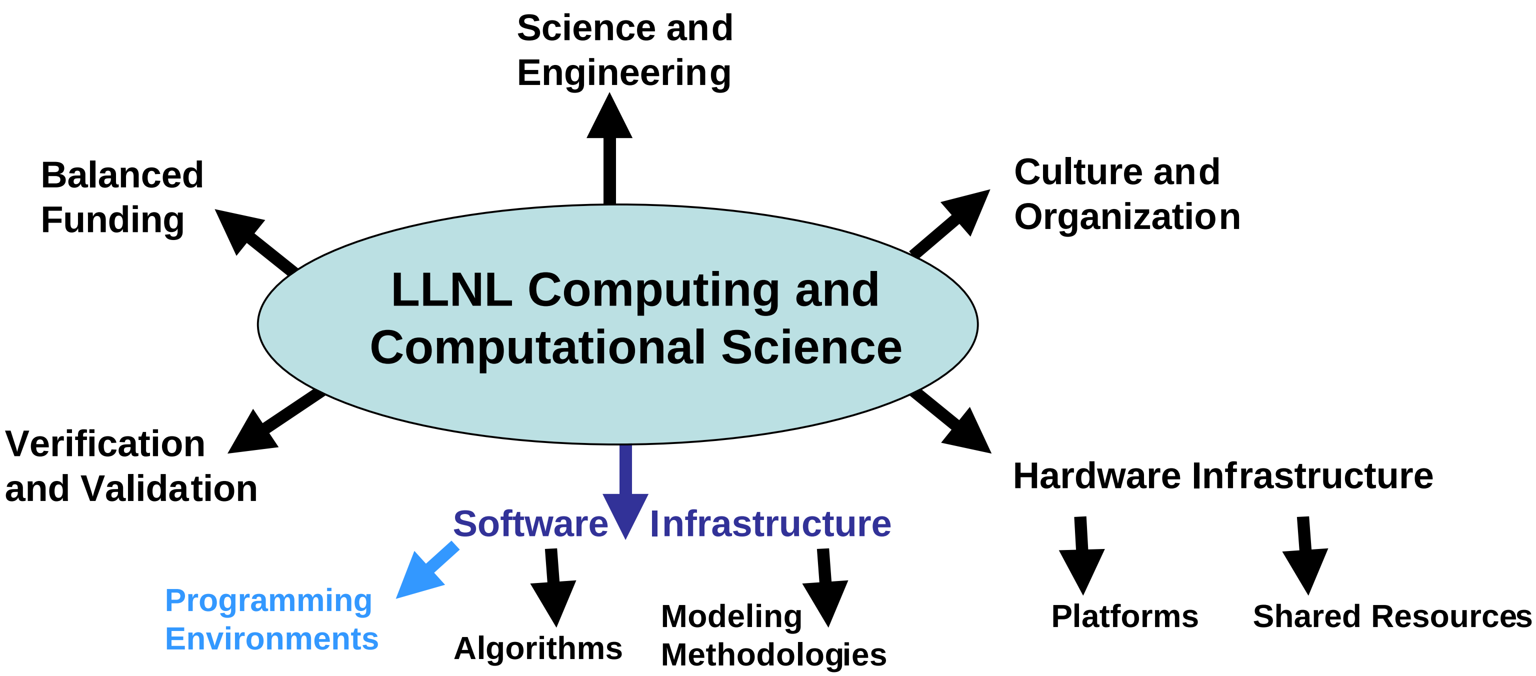
Finally, "the network is the computer"



Filling vendor gaps in shared-resource and other system software

- Past
  - Pioneered mass storage system software — PDP 10 Elephant system, UniTree system, High Performance Storage System (HPSS) (with IBM and other DOE labs)
  - Pioneered supercomputer OSs and compilers written in a higher level language
    - LTSS, NLTSS
    - Languages and compilers such as LRLtran (written in Fortran)
  - Pioneered many other shared and application-support tools
- Present
  - ASCI Pathforward, collaborations with universities and other labs
- Future
  - Fill the evolving vendor tool gaps alone, or in partnerships, or by other mechanisms such as open-source collaboration

Programming environments at LLNL

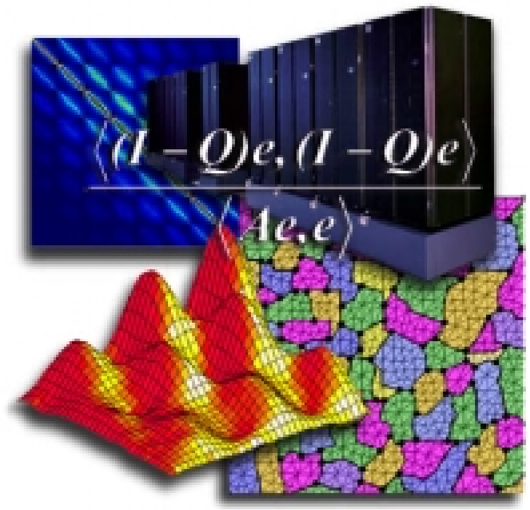


Programming environments change slowly over time

Programming environments (PEs) comprise languages, runtime environments, and tools to map algorithms to code, facilitate correctness, meet performance goals, and support code development

Building complex software has been, is, and will continue to be difficult

- PEs change slowly O (15+yrs)
- Software increases in size and complexity, and architectures change faster than improvements in PEs
- Architectural details are hidden to promote productivity and portability vs. exposed for performance optimization



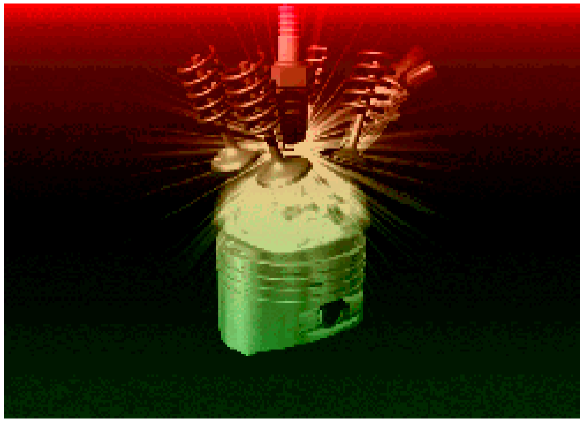
The CHI (Χ) Factor: the computer–human interface

PAST
Front-end console

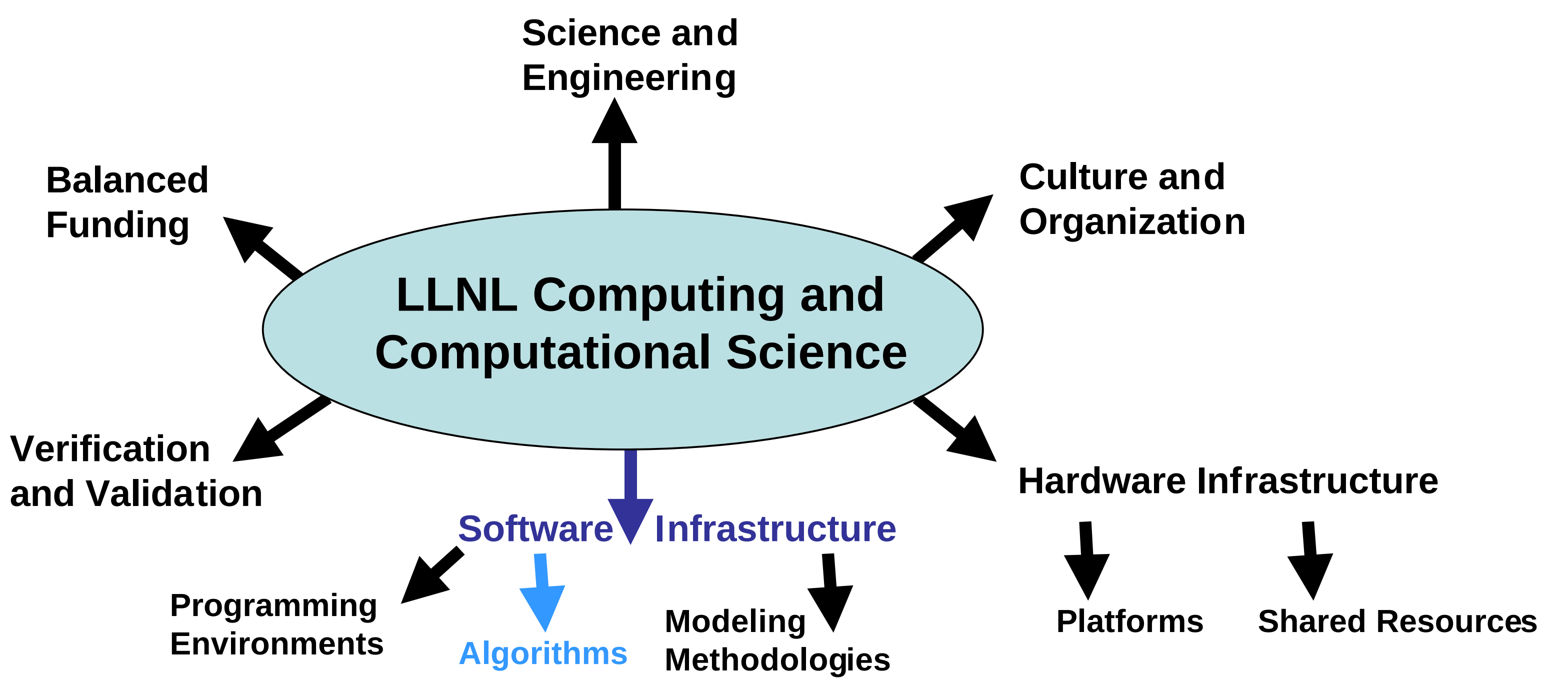
PRESENT
Keyboard and mouse

FUTURE
Holographic avatar

Image courtesy of
Holographic Studios
http://www.holostudios.com/#gall

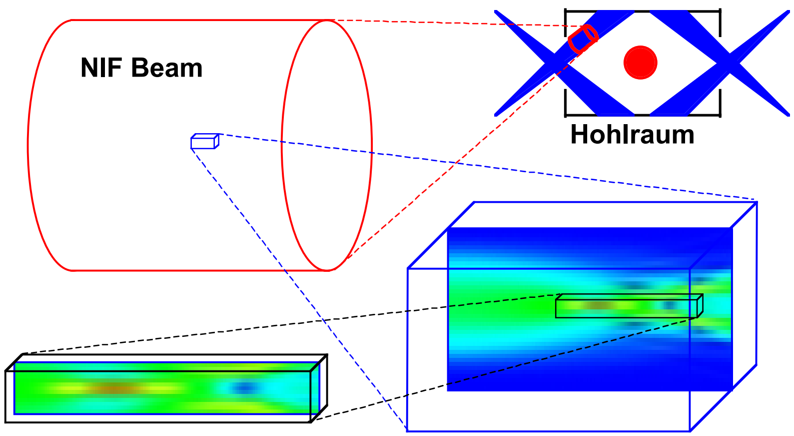


Algorithm development at LLNL



Algorithms advance as architectures evolve

- Algorithm advances must match architectural advances and change as architecture changes
- Past focus: to map science to computing and to optimize
- Present focus: to improve scalability
- Future focus: to develop better scalability, adaptivity, latency-hiding and deeper physical insights

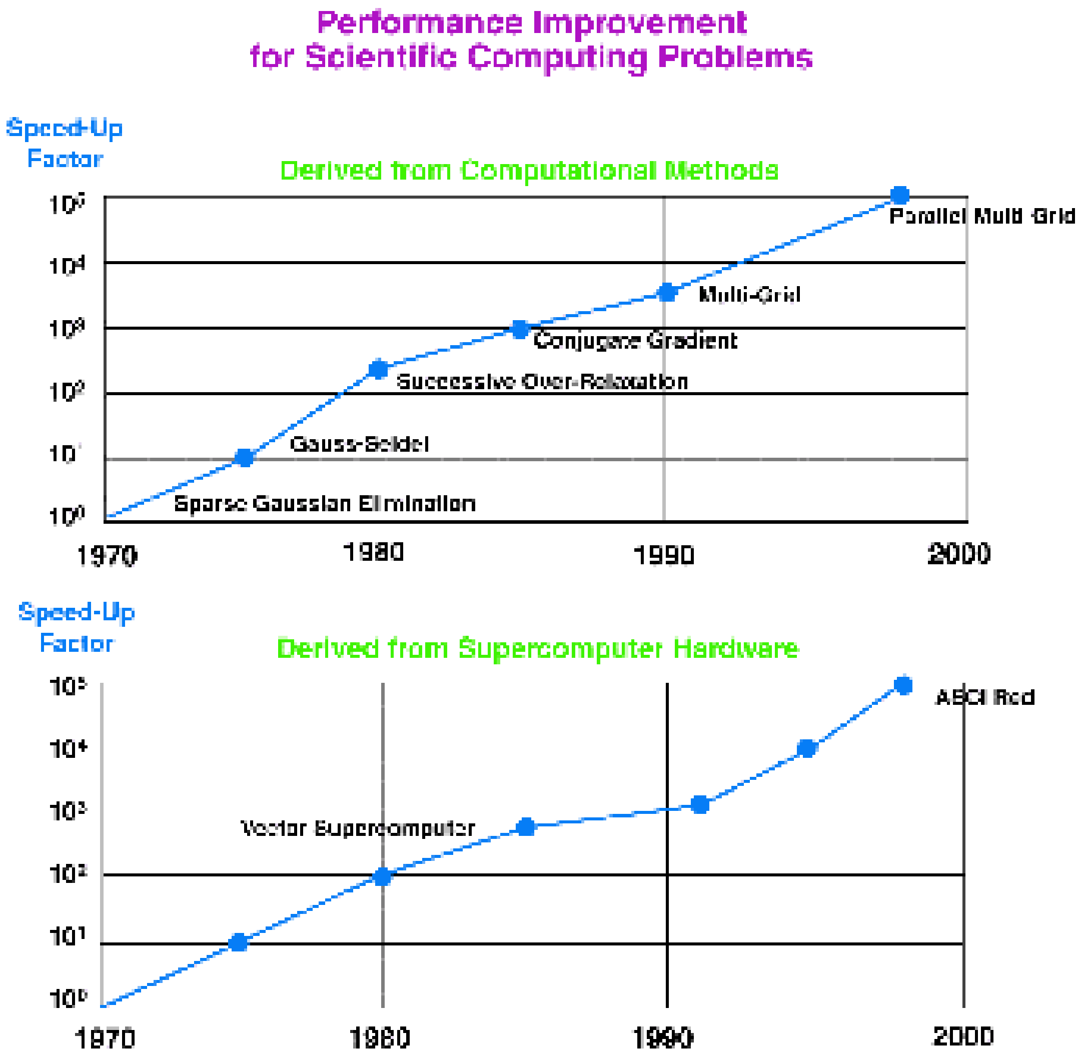


Adaptive Mesh Refinement (AMR) allows one to focus computational resources where they are most needed

An example of the importance of optimal algorithms

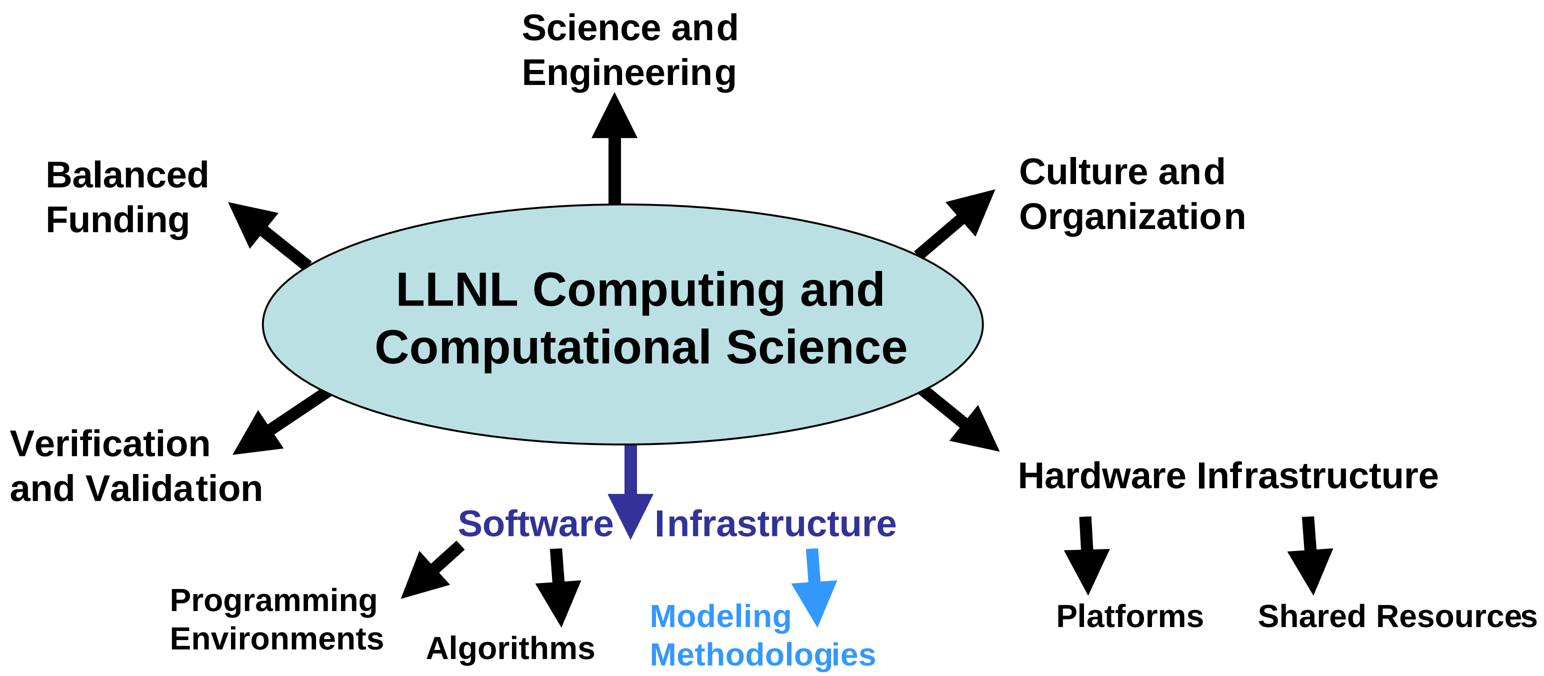
- **It is sometimes argued that increasing power from architecture diminishes the importance of algorithmic research, but the opposite holds**
- **The more powerful the computer, the *greater* the importance of optimality,**
  - **Suppose *Alg1* solves a problem in time CN^2, where N is the input size**
  - **Suppose *Alg2* solves the same problem in time CN**
  - **Suppose that the machine on which *Alg1* and *Alg2* run has 10,000 processors, that have been parallelized to run in constant time (compared to serial time), *Alg1* can run a problem 100X larger, whereas *Alg2* can run a problem 10,000X larger**
  - **In 2D 10,000X disappears fast; in 3D even faster**
- **Alternatively, filling the machine's memory, *Alg1* requires 100X time, whereas *Alg2* runs in constant time**
- **Large 10,000- processor machines are expensive, and optimal algorithms are the only algorithms that we can afford to run on them**

Algorithms are important contributors to total application performance



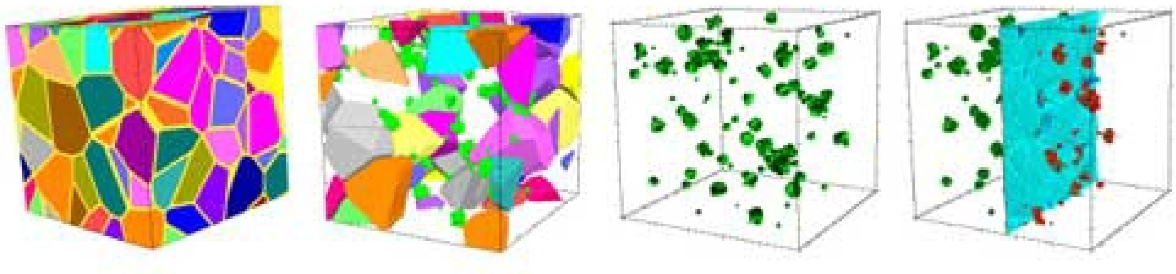
Total speedup:
algorithms x architecture

Modeling methodologies at LLNL



Modeling methodologies consider various factors

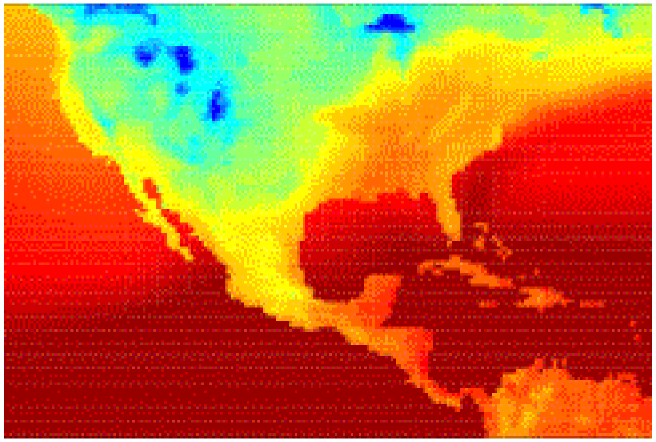
- Interacting dimensions of modeling methodologies
  - Model building
  - Performing a computational simulation experiment
- Model-building is difficult
  - It takes 1 to 5-ish years to build a model (models can last 20+ years)
  - What is needed and what is desired change over time
  - Developers need to adapt algorithms as computer hardware evolves



Modeling voids in materials is difficult work

Modeling methodologies are now tackling 3D meshes

- Performing simulation experiments
  - Problem setup
  - Run the model
  - Model output data management
  - Model output analysis and visualization



Atmospheric distribution simulation

In the future, there will be intuitive easy-to-use interfaces to assemble low- to middle-complexity models for routine model-building and simulation experiment running

Modeling data in relation to other data

What will we do with all the data?

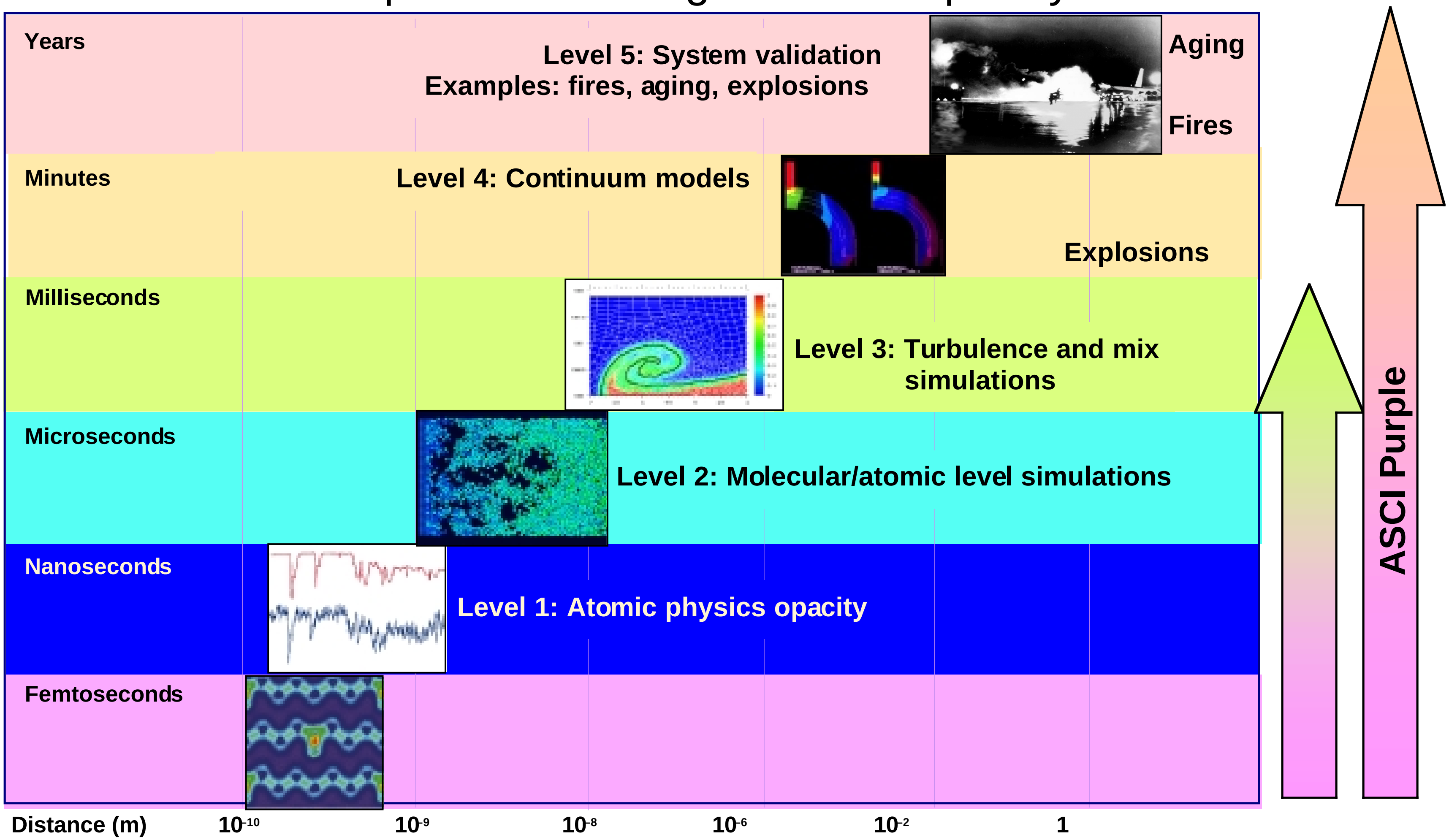
- Present: Estimates of information in the world
  - Library of congress 3 PB including all media
  - WWW 7 - 10 PB including accessible data bases such as genome and geo/climate data
  - Per year increases
    - Writing 200 TB
    - Movies 200 TB
    - Images 500 PB
    - Broadcasting 80 - 100 PB
    - Sound 60 TB
    - Telephony if digitized and captured 4 EB
- Future: In 2010, estimate 5 exabytes of hard disk shipped

In 2050, 100-zettaop system could be generating yottabytes/year

How will we organize and make sense of zettabytes of modeling and sensor network data?

ASCI simulation time scales span femtoseconds to years

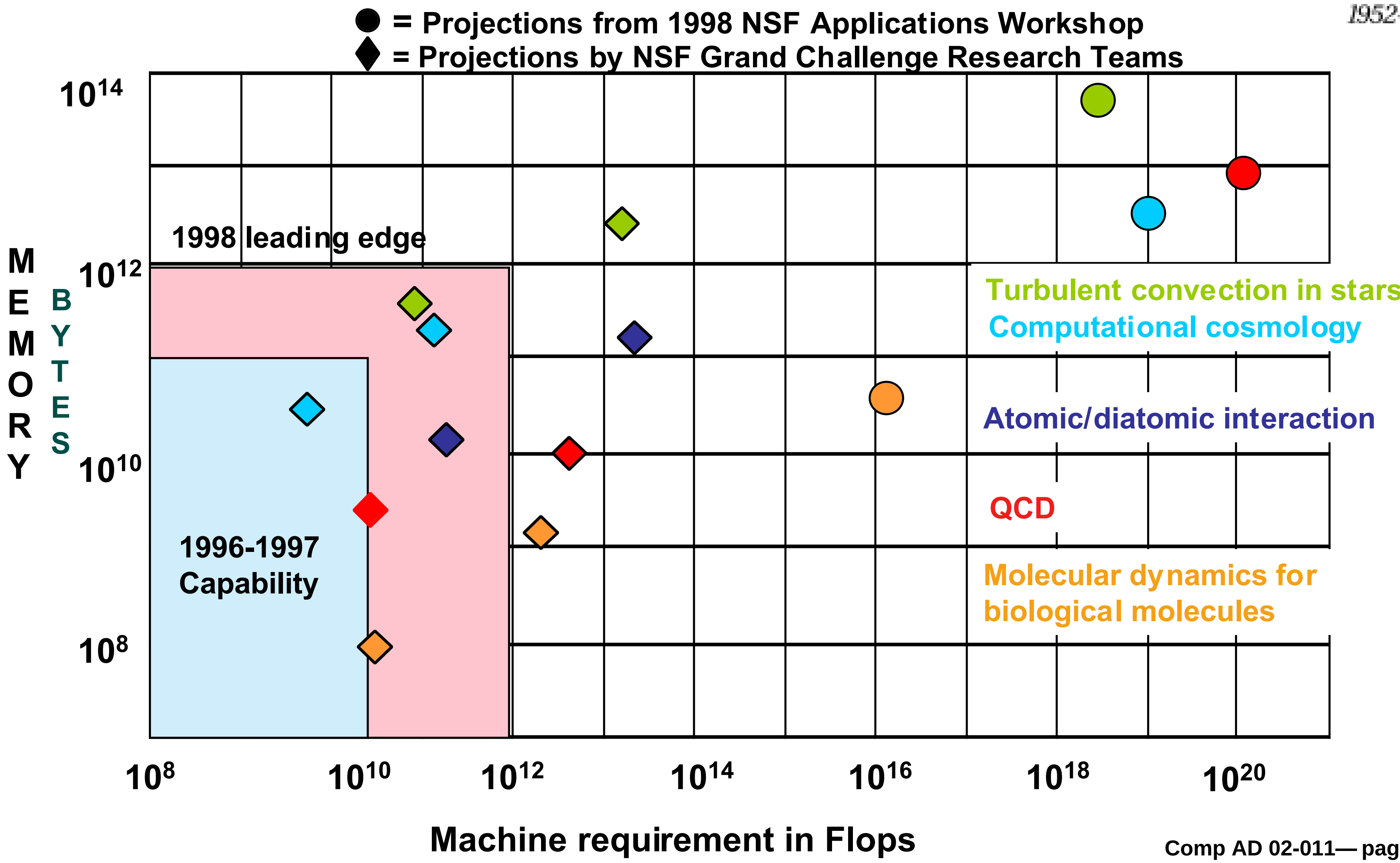
One example of increasing model complexity



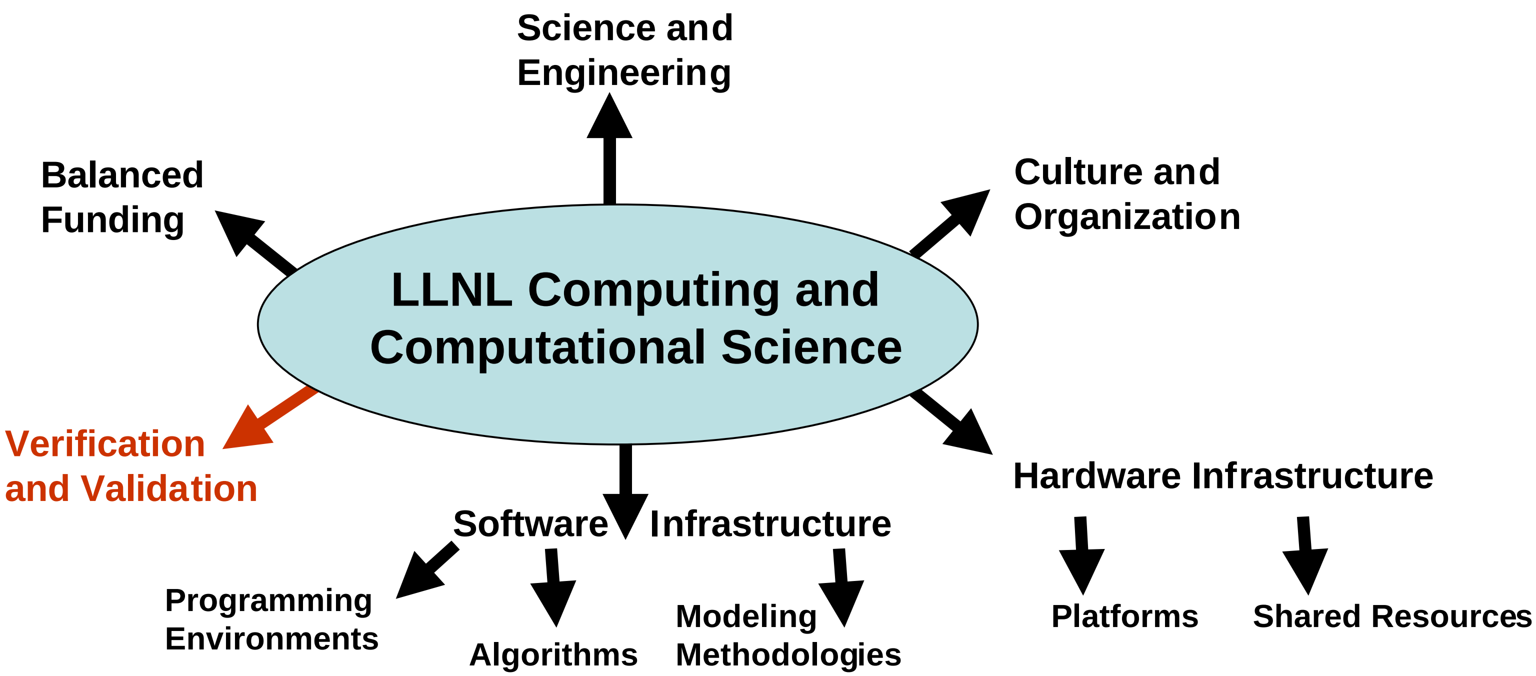
Exaflops only scratch the surface of potential for biological modeling

Computer	Number of atoms	Time scale	Biochemical processes
Gigaflop	<1,000	1 psec	Solvation small biochemicals Solvent phase small molecular reaction
Teraflop	10,000	100 psec	Structural relaxation modified DNA Enzyme catalyzed reaction-active site
Petaflop	100,000	10 nsec	Complete metabolic reaction process Protein-drug binding
Exaflop	1,000,000	1 µsec	Initial stages of protein folding Function of full DNA polymese reaction

Emerging computational opportunities: requirements for solution in 1 week



Verification and Validation

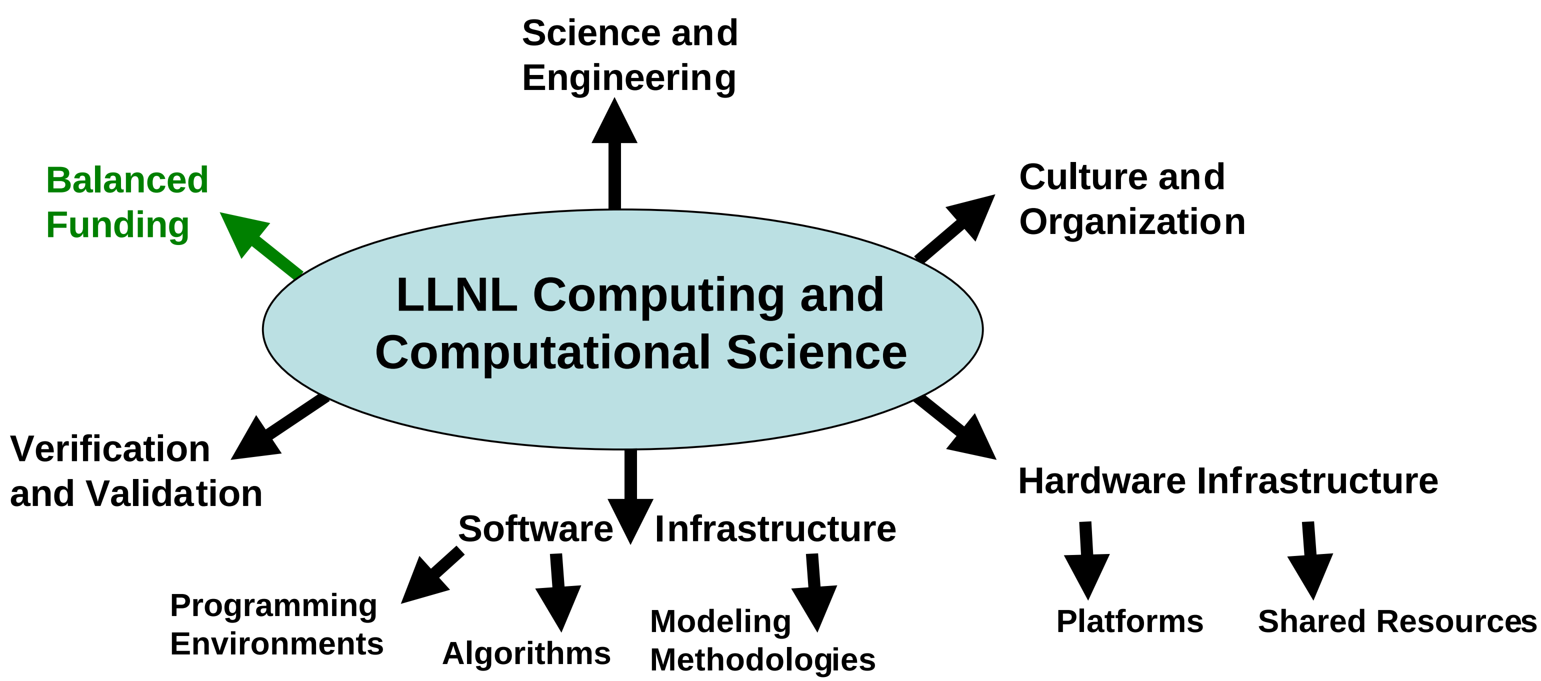


Verification and Validation are critical to simulation success

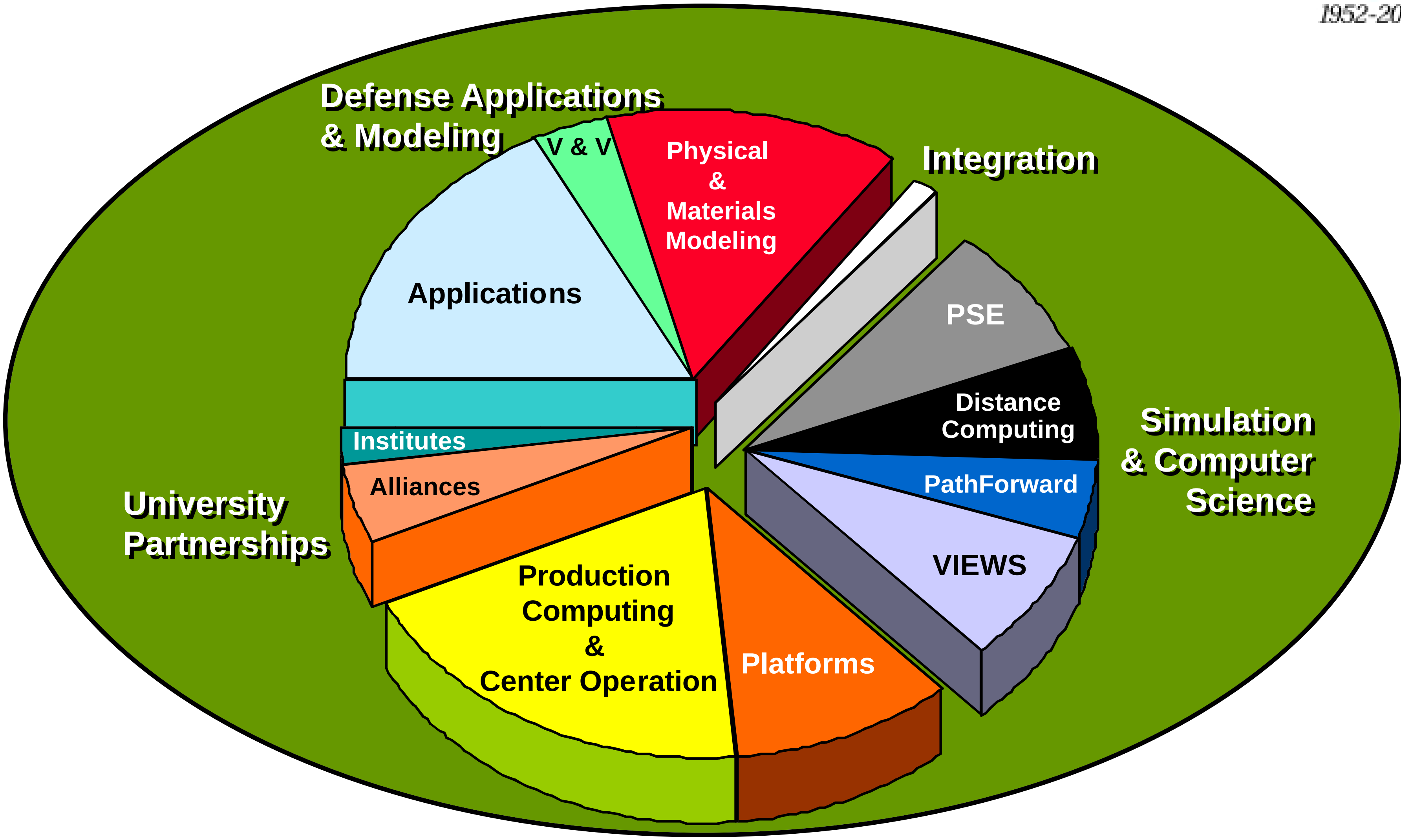
- **Verification** ensures that the software implementation is as intended, agrees with analytical solutions, uses modern software practices
- **Validation** evaluates the degree to which a computer model is an accurate representation of the real world via comparisons with data
- Past
  - V&V was less formal, codes smaller, less coupled physics, and nuclear tests ensured devices were modeled adequately
  - Model input and parameters could be carefully tuned to match experimental results
- Present and future
  - V&V has become a major scientific and computer science focus in the computational simulation community, at LLNL and in the world
  - Because models are built from modules representing many scientific domains, many areas of experimentation are required

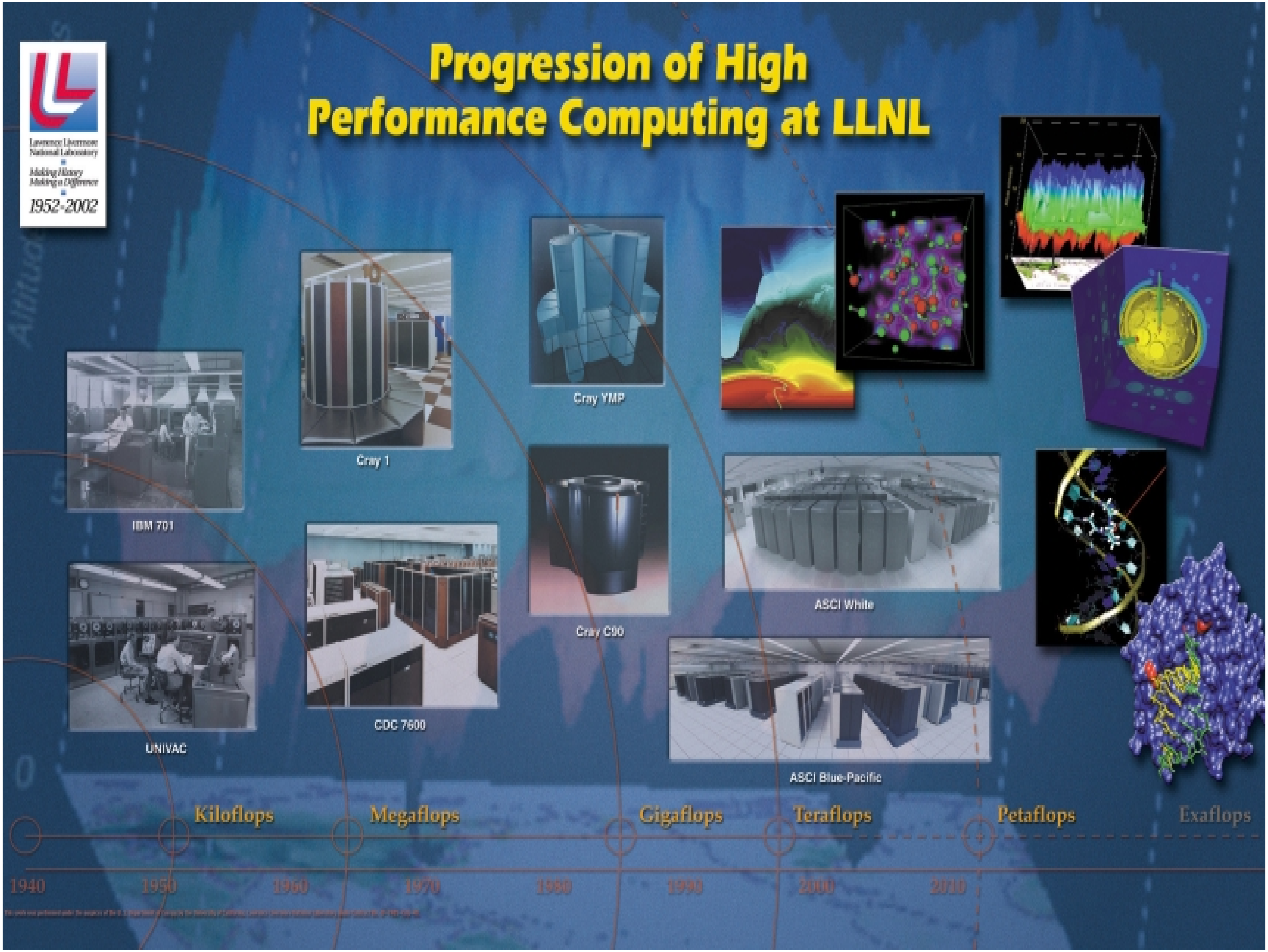
Validation requires a range of experiments for complex models

Balanced funding



ASCI example of the elements of balanced funding




Progression of High Performance Computing at LLNL
Lawrence Livermore National Laboratory
Making History Making a Difference
1952-2002
IBM 701
UNIVAC
Cray 1
CDC 7600
Cray YMP
Cray C90
ASCI White
ASCI Blue-Pacific
Kiloflops
Megaflops
Gigaflops
Teraflops
Petaflops
Exaflops
1940
1950
1960
1970
1980
1990
2000
2010

Summary

- **Experience at LLNL underscores that the key issues important 50 years ago are important today, and will be important into the future**
  - **Only the details will change**
- **Keep focus on**
  - **Pushing the frontiers of HPC balanced hardware and software infrastructure**
  - **Strong personal and organizational commitment to the computational science research program**
    - **Push hard on what can be modeled, and on model accuracy**
    - **Support scientific and modeling culture and organization**
  - **Verification and validation methodologies**
  - **Adequate, balanced funding**
- **Strive to achieve balance within and across computing dimensions**